

지역 기상 정보를 활용한 단기 전력 수요 예측 모델

A New Prediction Model for Power Consumption with Local Weather Information

탁해성, 김태용, 조환규, 김희제
부산대학교 전기전자컴퓨터공학과

Haesung Tak(tok33@pusan.ac.kr), Taeyong Kim(tayong22@pusan.ac.kr),
Hwan-Gue Cho(hgcho@pusan.ac.kr), Heeje Kim(heeje@pusan.ac.kr)

요약

많은 정보가 데이터로 저장되면서, 데이터를 분석하거나 특수 상황을 예측하기 위한 연구들이 진행되고 있다. 특히, 전력 데이터의 경우 환경적 요인에 의한 예측 연구 및 신재생 에너지를 활용하는 스마트그리드와 마이크로그리드 연구들이 진행 중이다. 본 논문에서는 전력 데이터의 예측을 위해 주변 환경에서 나타나는 데이터를 활용하고자 한다. 이때, 단순 기상 데이터가 아닌 이전 시간에 따른 여러 인자를 반영하여 데이터 예측이 올바르게 이루어지는지를 검증하고자 한다. 검증 과정에서는 유사 기온을 가지는 전력 데이터 선별 예측 결과와 전력 데이터의 길이에 따른 전력 수요 예측 결과를 비교군으로 두고, 기상 정보를 추가 활용하였을 때의 전력 수요 예측 결과를 비교 분석한다. 실험 결과를 통해 기상 정보를 이용할 경우 평균 15% 이내의 최대 오차율 감소 효과를 확인할 수 있다.

■ 중심어 : | 전력 예측 | 서포트 벡터 회귀 | 단기 예측 |

Abstract

Much of the information is stored as data, research has been activated for analyzing the data and predicting the special circumstances. In the case of power data, the studies, such as research of renewable energy utilization, power prediction depending on site characteristics, smart grid, and micro-grid, is actively in progress. In this paper, we propose a power prediction model using the substation environment data. In this case, we try to verify the power prediction result to reflect the multiple arguments on the power and weather data, rather than a simple power data. The validation process is the effect of multiple factors compared to other two methods, one of power prediction result considering power data and the other result using power pattern data that have been made in the similar weather data. Our system shows that it can achieve max prediction error of less than 15%.

■ keyword : | Power Prediction | Support Vector Regression | Short-term Prediction |

1. 연구동기

일상에서 흔히 측정되는 정보들이 전산화되기 시작하면서, 데이터의 추이를 예측하고자 하는 시도들이 이

루어지고 있다. 전력 발전소에서 생산되어 사용되는 데이터도 예외가 될 수 없으며, 스마트그리드와 마이크로그리드 개념이 등장하면서 각 지역의 시간별 전력 수요 및 최대 부하를 예측하는 것이 큰 쟁점이 되고 있다[1].

* 이 논문은 2015년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 기초연구실육성사업 연구임

(NRF-2015R1A4A1041584)

접수일자 : 2016년 09월 21일

수정일자 : 2016년 11월 17일

심사완료일 : 2016년 11월 17일

교신저자 : 조환규, e-mail : hgcho@pusan.ac.kr

스마트그리드의 작은 개념인 마이크로그리드는 많은 신재생에너지 자원과 전력 저장 도구와 같은 전력 제어 모듈 등이 연결되어 동작한다. 전력 자원과 제어 모듈들은 상호 동작이 효율적으로 이루어져야 하며, 그렇지 못할 경우 전력망의 안정성에 악영향을 주게 된다. 따라서 다양하게 변화하는 환경 및 전력 수요를 예측하고 대응할 수 있는 시스템의 필요하다[2][3].

본 논문에서는 전력 및 기상 데이터 특성을 이용하여 전력 예측 가능성을 검증하고자 한다. 데이터를 구성하기 위해 시간별 전력 / 기상 데이터를 활용하고자 하며, 실험 방법에 따라 데이터를 분류하여 유사 연구들과의 차이를 분석하고자 한다. 전력 예측을 위해 기계 학습 기법인 SVM(Support Vector Machine)에 회귀 분석 함수를 적용한 SVR(Support Vector Regression)을 이용하였다[4]. 본 연구에서 사용한 SVR의 비교 실험을 위해 유사 기온 데이터를 가지는 전력 패턴 데이터와 전체 데이터에 대한 비교 실험을 통해, 제안 기법의 예측 효율성을 검증한다.

II. 관련 연구

전력 수요 예측 알고리즘은 기존 데이터를 활용하여 현재 데이터를 추정하는 통계 기법과 입출력 사이의 관계를 추정하는 인공지능 기법이 활발하게 연구되고 있다. 수요 예측 통계 기법으로는 데이터 군집에 대한 회귀 분석, 시계열 기반의 신뢰 구간 추론 등이 있으며, 전력 수요에 영향을 주는 다양한 인자들을 적용하여 비선형적 특성을 분석하는 SVM이 있다. SVM은 기계 학습 분야 중 하나로 패턴 인식, 자료 분석을 위한 학습 모델이며, 분류와 회귀 분석에 효율적이다. SVM 알고리즘은 주어진 데이터 집합을 바탕으로 하여 새로운 데이터에 대한 비확률적 선형 분류 모델을 이용한다. SVM은 선형 분류와 더불어 비선형 분류에서도 사용될 수 있다. 인공지능 기반의 수요 예측 기법으로는 신경망 네트워크 기법 및 퍼지 이론 등이 있으며 하드웨어의 발달로 인한 딥러닝 기법도 고려되고 있다. 최근에는 기존 전력 수요 예측 모델의 개선을 위해 한 가지 이상의 시스템을 혼합하는 하이브리드 모델이 제시되고 있다.

퍼지 인공지능망 기법과 웨이블릿 퍼지 인공지능망 기법을 혼합한 하이브리드 모델을 이용한 수요예측 연구 및 SVD(Singular Value Decomposition)기반 지수 평활 기법에 다양한 지수 평활 가중치를 적용하는 연구를 비롯한 다양한 기법을 적용하는 연구가 진행되고 있다[5-7].

1. 해외 주요 전력수요예측 연구 및 개발 현황

PJM은 미국의 펜실베이니아와 뉴저지, 메릴랜드 등 13개 주와 워싱턴 DC의 전력을 공급하는 독립 계통 운영 업체이다. PJM은 전력수급과 관련된 다양한 변수들을 고려하여 가상 시나리오를 준비하고, 비상사태에 대비하고자 한다. 이때, 전력이 영향을 미칠 수 있는 요인으로 이상기온, 비상상황, 상정사고 등을 고려한다. 또한, 단기 전력 수요 예측을 위해 신경 회로망(Neural Network) 모델과 유사 전력 패턴 모델을 사용하여 전력 수요 예측 및 가격 결정에 활용하고 있다. 이때 활용되는 인자로는 평일 및 주말에 따른 요일 분류, 공휴일 및 기타 특수 상황이 발생한 날짜, 기상 관측 자료, 전력 수요 조정량 등을 고려하여 예측을 시행하고 있다[8, 9]. PJM은 기상상황에 대하여 크게 'Normal', 'Extreme', 'Mild' 3가지로 분류하여 전력 수요 예측에 활용하고 있으며, 연간 평균 오차는 2% 범위를 유지하고 있다. [그림 1]은 PJM의 전력 수요 예측 프로세스에 대한 그림으로, 수요예측이 완료되어도 수요예측 결과는 관련 인자와 실제 기상에 맞추어 재조정 과정을 거쳐 정밀한 예측 결과를 제공하는 것을 확인할 수 있다.

이와 달리 미국 캘리포니아 주에서 독립 계통 운영 회사로 알려진 CAISO의 경우, 캘리포니아의 송전망 관리 및 전력 시장 활성화 등 다양한 역할을 수행하고 있다. CAISO는 기후적 특징을 기준으로 10개의 기후지역을 분류하고, 지역별 수요예측자료가 전체 예측 자료로 취합되는 상향식 모델을 기반으로 하고 있다. 수요 예측 모델은 신경 회로망 모델과 ARIMA를 기반의 ALFS(Automated Load Forecasting System) 프로그램을 활용하고 있으며, 기상자료는 24개의 기상관측지점의 9일 예보자료로써, 기온, 이슬점, 풍속, 조도 등을 사용하고 매시간별로 기상 실측자료를 업데이트한다. 하루 전 수요예측은 전일 오전 9시에 시행하고 있다[10].

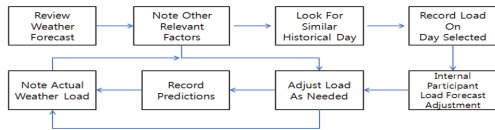


그림 1. PJM의 전력 수요 예측 프로세스

2. 국내 수요예측 연구 및 개발 현황

한국 전력 거래소는 단기 전력 수요 예측 프로그램(KSLF : KPX Short-Term Load Forecaster)을 개발하여 수요예측업무에 사용하고 있다. KSLF는 과거 전력 수요 패턴 분석을 통해 기온과 전력 수요 특성을 고려한 수요 예측을 시행한다. 일반적으로 평일과 주말은 기상 요인을 반영한 지수 평활 기법을 사용하며, 공휴일을 포함한 특수일은 퍼지 선형 회귀 분석 기법을 활용하여 기온 변화를 고려하지 않은 수요 예측 결과를 계산한다[2][11][12].

한국 전력 거래소에서 수요 예측에 사용하는 데이터는 시간 단위 전력 데이터와 일일 최고/최저 기온, 특수 상황에 대한 조사 조업률이다. 지역별 전력 수요가 아닌 전국 단위의 전력 수요를 예측하기 때문에, 전국 8개 도시 기온을 이용하여 전국 대표기온을 추정 후 수요 예측을 시행한다. 이외에도 통계 기법을 적용한 전력 수요 예측 연구가 진행되고 있다. 연평균 전력 사용량을 고려한 연간 전력 수요에 대한 예측을 비롯하여 일별 기온 특성 및 기계 학습을 이용한 수요 예측 연구가 그 예이다[13][14].

III. 전력 데이터의 특성 비교

1. 지역 특성에 따른 전력 수요 패턴 분석

전력 수요 예측 모델을 제안하기에 앞서, 부산 지역 내 각 변전소의 전력 수요 패턴을 분석하였다. 이는 변전소의 지역 및 기간 특성에 따라 전력 수요 변화를 분류하기 위함이다. 변전소의 위치 특성에 따라 주로 나타나는 패턴은 2가지로, 일부 지역에 따라 전력 수요 패턴이 다르거나 수요가 작아 패턴을 구분 할 수 없는 변전소가 존재한다. 하지만 특정 지역에서 수요 패턴이 다르게 나타나는 경우에도 주거지 및 공단에 따른 특성

이 나타나므로, 본 논문에서는 부산에 위치한 변전소에서 나타나는 전력 패턴을 공단, 주거, 특이 지역으로 분류하였다. 특이 지역의 경우 일정한 전력 수요 패턴이 나타나지 않는 지역을 의미한다. 공단 및 주거 지역 전력 수요 패턴은 [그림 2], [그림 3]과 같다.

[그림 2]는 공단 지역으로 분류된 녹산 지역 변전소의 전력 수요 패턴을 시각화한 결과를 나타낸다. [그림 2]의 전력 패턴을 일일 단위로 비교하였을 때, 공단 지역에서는 주말과 주중 전력 수요 패턴이 다르게 나타남을 알 수 있다. 계절 특성에 무관하게 주중 전력 수요가 주말보다 높은 것을 알 수 있으며, 공휴일의 경우를 제외하였을 때 일요일의 전력 수요가 가장 최소로 나타나는 것을 확인할 수 있다. 일일 최대 피크 시간대는 주로 점심시간 이후인 오후 2~5시경이며, 일일 최소 전력 수요 시간은 공장 가동되기 직전인 오전 5~7시 사이로 나타난다. 또한, 점심시간인 오후 1시경에 전력 수요가 급감하는 것을 확인할 수 있다.



그림 2. 지역 특성에 따른 공단 분류 지역(녹산)의 전력 수요 패턴. (a) 주말 전력 수요, (b) 주중 전력 수요, (c) 일일 최대 전력, (d) 일일 최소 전력, (e) 점심 시간의 전력 수요 [x축 : 시간 ($|T_i| = 168$), y축 : 전력 수요]

[그림 3]은 주거 지역으로 분류된 동래 변전소의 전력 수요 패턴 시각화 결과를 나타낸 것이다. 주거 지역은 주말 및 주중의 전력 수요가 거의 같으며 평균적으로 저녁 시간대에 수요 크다. 전력 수요 패턴의 경우, 주말에 일부 차이가 존재하지만 수요 패턴 모형은 일정하게 나타나는 것을 알 수 있다. 또한, 저녁 시간대인 오후 8시에 최대 전력 수요가 나타나며, 최소 전력 수요는 공장 지역과 유사한 오전 5~7시 사이로 나타나는 것을 확인할 수 있다.

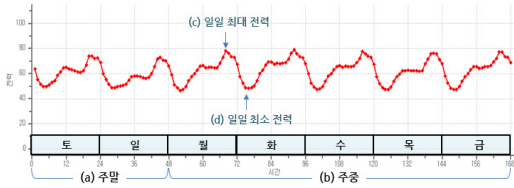


그림 3. 지역 특성에 따른 주거 분류 지역(동래)의 전력 수요 패턴. (a) 주말 전력 수요, (b) 주중 전력 수요, (c) 일일 최대 전력, (d) 일일 최소 전력 [x축 : 시간, y축 : 전력 수요]

2. 계절 특성에 따른 전력 수요 패턴 분석

본 논문에서는 지역적 전력 수요 패턴 이외에도 계절에 따른 전력 수요 패턴을 비교 분석하였으며, 이는 [표 1], [표 2]의 분석 결과를 통해 알 수 있다.

[표 1]은 다대 지역의 기온 정보에 대한 통계 분석 결과를 나타낸 것으로, 사하 방재 관측기기의 2011년부터 2014년까지의 기온 데이터 분석 결과를 나타낸다. 방재 관측기기의 경우 기온, 풍속, 풍향, 강우량 데이터가 저장되며, 전력 수요에 영향을 미치는 주요 요인으로 기온, 강우량을 사용가능하나 2014년도의 전반적인 강우량 수치가 낮게 나타나 기온 데이터만을 활용한다. 기온 데이터 분석 결과 계절별 차이를 확연하게 알 수 있다. [표 2]는 다대 변전소의 월별 전력 수요를 분석한 결과로 평균 기온이 가장 높은 8월과 평균 기온이 가장 낮은 1월이 최고 전력 수요가 나타나는 것을 알 수 있으며, 평균 전력 수요가 높은 계절은 여름과 겨울임을 알 수 있다.

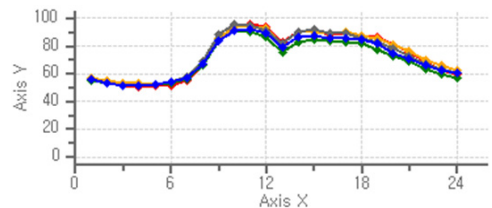
[그림 4-7]은 계절에 따른 다대 지역의 전력 및 기온 데이터를 나타낸다. 다대 지역은 지역 특성에 의한 분류 중 공단 지역의 특성이 강하다. 각 계절의 대표 전력 및 기온 데이터는 2014년을 기준으로 1, 4, 7, 10월의 각 4주의 화요일 데이터를 수집한 것이다. 각 대표 데이터의 데이터 추이 변화를 통해 기온에 따른 전력 사용량 변화가 나타나는 것을 확인할 수 있다. 각 그림을 비교하였을 때, 전력 수요의 경우에는 겨울과 여름, 봄과 가을이 각각 유사하게 나타나며, 기온 데이터는 각 대표 데이터에서 다르게 나타난다.

표 1. 다대 주변 기상 방재 관측소인 사하 관측소의 월별 기온 분석 결과 [기온 단위 : °C]

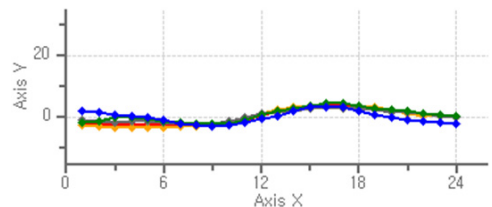
월	평균 기온	최고 기온	최저 기온	분산
1	1.68	15.30	-11.50	17.99
2	3.88	18.00	-10.80	19.80
3	8.31	22.40	-3.10	19.53
4	13.14	26.90	2.50	17.36
5	18.41	31.20	8.30	15.74
6	21.56	30.10	14.40	7.97
7	25.96	35.20	19.00	9.03
8	26.34	37.10	18.70	10.97
9	22.12	32.80	11.90	10.95
10	16.72	27.50	3.90	14.62
11	10.76	23.30	-0.90	23.18
12	3.15	13.40	-6.80	16.81

표 2. 다대 변전소의 월별 전력 수요 분석 결과 [전력 단위 : kWh]

월	평균 전력	최고 전력	최저 전력	분산
1	93.05	130.10	41.60	295.44
2	89.91	127.70	45.10	288.18
3	86.34	121.80	46.00	148.56
4	80.98	103.30	45.50	116.61
5	77.83	114.70	39.10	143.38
6	77.44	103.30	37.50	158.49
7	87.25	128.20	38.30	319.99
8	86.76	133.10	38.60	368.23
9	76.75	128.50	36.60	256.86
10	75.50	129.70	36.00	146.76
11	81.15	109.40	42.20	162.43
12	92.40	130.70	50.80	208.39

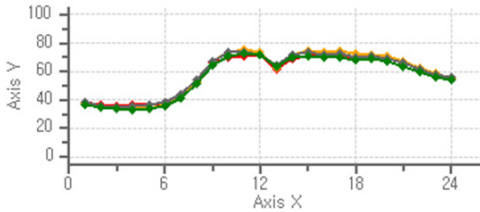


(a) 전력 수요 시계열 데이터 [x축 : 시간, y축 : 전력 수요]

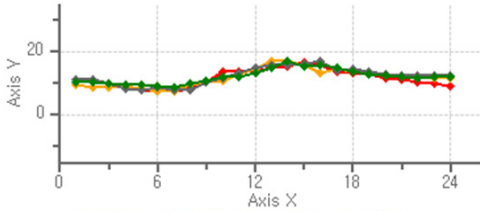


(b) 기온 시계열 데이터 [x축 : 시간, y축 : 기온]

그림 4. 다대 지역 변전소의 겨울철 대표 전력 및 기온 시계열 데이터 예시. (a) 전력 수요 / (b) 기온 시계열

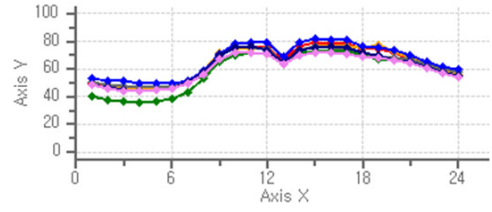


(a) 전력 수요 시계열 데이터 [x축 : 시간, y축 : 전력 수요]

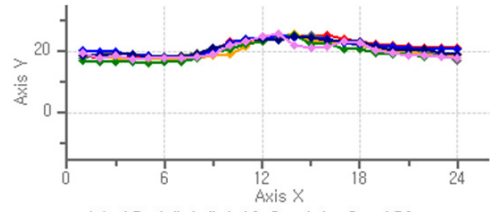


(b) 기온 시계열 데이터 [x축 : 시간, y축 : 기온]

그림 5. 다대 지역 변전소의 봄철 대표 전력 및 기온 시계열 데이터 예시. (a) 전력 수요 / (b) 기온 시계열

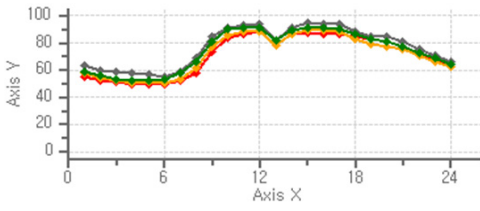


(a) 전력 수요 시계열 데이터 [x축 : 시간, y축 : 전력 수요]

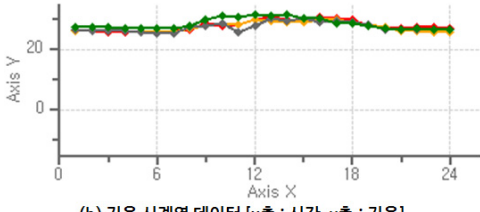


(b) 기온 시계열 데이터 [x축 : 시간, y축 : 기온]

그림 7. 다대 지역 변전소의 가을철 대표 전력 및 기온 시계열 데이터 예시. (a) 전력 수요 / (b) 기온 시계열



(a) 전력 수요 시계열 데이터 [x축 : 시간, y축 : 전력 수요]



(b) 기온 시계열 데이터 [x축 : 시간, y축 : 기온]

그림 6. 다대 지역 변전소의 여름철 대표 전력 및 기온 시계열 데이터 예시. (a) 전력 수요 / (b) 기온 시계열

[그림 4], [그림 6]의 전력 수요가 유사한 것을 통해 기온 차이가 나타나더라도 냉난방기 사용으로 인해 전력 수요가 높음을 알 수 있으며, [그림 5], [그림 7]에서 전력 수요수치가 유사한 것을 알 수 있다.

각 계절 특성에도 전력 수요 패턴 모양은 달라지지 않는 것을 알 수 있다. 이는 기온 변화에 대한 전력 수요 민감도가 인구의 활동 시간이 전력 수요에 영향을 주는 것을 의미한다.

III. 기후 데이터를 이용한 전력 예측 과정

1. 전력 수요 예측 과정

본 논문에서 전력 및 기상 데이터를 고려하여 전력 예측을 하기 위한 과정으로 전력 수요 및 기상 정보에 대한 파일 호출 및 병합 과정이 이루어지며, 지역에 따른 학습 및 평가 데이터 선정, SVR을 이용한 데이터 학습 및 예측, 결과 비교 평가 순서로 예측이 진행된다. [그림 8]은 전력 수요 예측 프로세스를 나타내며, 전력 및 기상 학습 데이터의 벡터 변화 및 매개변수 설정, 예측 기법을 달리하여 SVR 사용에 따른 예측 결과를 검증하고자 한다.

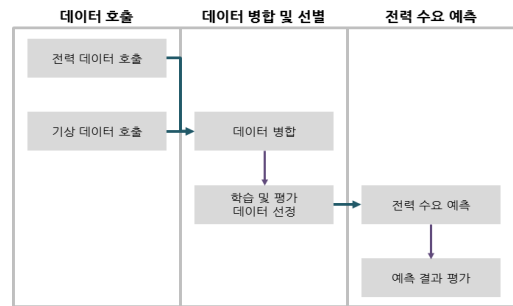


그림 8. 본 논문에서 제안하는 환경 변수를 이용하는 전력 수요 예측 프로세스

2. 전력 및 기상 데이터 병합

전력 데이터에 영향을 주는 요인으로 계절 및 휴일과 같은 상황적인 특징과 인구수 증가, 지역 특성과 같은 환경적 특성이 존재한다. 본 연구에서는 같은 전력 데이터 및 기상 데이터의 시간적 유사성을 기준으로 기상 및 전력 데이터에 따른 특성을 분석한다. 이를 기준으로 단일 전력 데이터에 대한 수요 예측 및 기상 정보를 추가로 사용한 수요 예측을 시행하여, 기상 정보를 전력 수요에 이용하였을 때의 비교 실험을 진행하고자 한다.

3. 학습 및 평가 데이터 선정

데이터 분류가 수행되고 나면, 예측 실험을 위해 학습 및 평가 데이터를 선정한다. 학습 및 예측 범위를 선정하기 위해 전력 및 기상 정보 데이터의 중첩 범위를 분석하였다. 전력 및 기상 정보 데이터에 대한 데이터 발생 기간을 분석한 결과, 2011년부터 2014년까지의 데이터가 중첩되는 기간으로 나타났다. 이에 따라, 2011년부터 2013년까지의 데이터를 학습에 이용하고, 2014년도 데이터를 평가 데이터로 활용하여 예측 실험을 진행하였다. 실험 데이터의 특성은 [표 4-6]을 통해 설명하고자 한다.

4. SVR을 이용한 전력 수요 예측

전력 수요 예측을 위해 전력 및 기상 데이터의 시간적 특성 및 이전 시간대의 데이터를 함께 사용하여 유사한 데이터에 대한 회귀 분석 결과를 예측에 활용한다. SVM은 실증적인 위험을 최소화하기 위해 사용되는 기법으로 기존 예측 기법들보다 학습 및 예측 결과의 정확도가 높게 나타나며, 비선형 데이터에 대한 학습도 가능하다. 회귀분석을 통한 전력 예측 연구의 경우 데이터 선별 및 회귀 함수 추정을 위한 계산 과정이 실시간으로 진행되기 어려워, 예측 결과를 도출하는데 많은 시간이 소요된다.

전력 데이터는 변전소별로 일정한 수요 모형이 존재하기 때문에 일반 SVM을 활용하는 것보다는 회귀 함수를 사용하는 것이 좋다. 본 논문에서는 SVM 성능 향상을 위해 회귀 함수를 적용한 SVR을 이용하고자 한다 [4]. SVR은 기본 모형을 만들기 위한 회귀 함수를 구성

하고, 학습 데이터를 통해 회귀 함수를 근사한다. 아래 식 1은 SVR에서 사용하는 회귀 함수의 직선식이다.

$$f(x_i) = \omega \cdot x_i + b \quad (1)$$

본 논문은 SVR을 활용하여 기존 연구의 예측 결과를 비교 분석하고자 한다. 회귀 함수 추정을 위해 입력되는 매개변수 정보는 [표 3]과 같다. [표 3]은 전력 수요 예측을 위해 주어지는 정보를 나타낸 것으로, 예측하고자 하는 시간 및 지역 정보, 예측 시간 이전 k 길이의 전력 및 기온 패턴을 이용하여 예상 전력 수요를 계산한다.

표 3. 전력 예측을 위해 주어지는 환경 변수 정보

환경 변수	데이터 구분
T_i	예측 시간
z_i	기상 관측 장비 지역 정보
$\langle p_i^{(k)} \rangle$	예측 시간부터 k 시간 이전까지의 전력 벡터
$\langle t_i^{(k)} \rangle$	예측 시간부터 k 시간 이전까지의 기온 벡터

제안하는 예측 모델은 식 1의 x_i 에 대한 값을 [표 1]의 벡터 형태로 전달하여 식 2와 같은 예측 함수 형태로 활용한다.

$$P_i = \text{predict}(T_i, z_i, \langle p_i^{(k)} \rangle, \langle t_i^{(k)} \rangle) \quad (2)$$

전력 수요 예측을 진행할 때, 전력 및 기온 패턴은 예측 시간으로부터 k 시간 이전까지의 관측된 부분 시계열 데이터를 활용한다. 본 연구에서는 단일 전력 데이터를 이용한 전력 수요 예측의 한계를 극복하기 위해 전력 수요 패턴 분석에서 나타나는 기상 정보를 반영한다. 공휴일과 같은 특수 상황의 경우 학습 데이터가 충분할 때, 예측 결과가 실제 데이터와 유사한 것을 바탕으로 예측을 위한 인자로 활용하지 않았다. 실험에 사용되는 기상 및 전력 데이터를 공동으로 활용하여 활용 가능한 정보가 2011년에서 2014년으로 제한되어 온도를 제외한 기후 정보는 활용하지 않았다. 이는 2011년부터 2014년까지 부산의 강수량과 천재지변이 다양하게 일어나지 않아, 변인으로 작용하지 않기 때문이다.

IV. 다중 특성에 따른 전력량 예측 실험

1. 실험 데이터

실험에 사용된 데이터는 부산 소재 변전소 및 기상 관측 기기에서 발생하는 데이터를 활용한다. [표 4-6]은 기상 관측소 및 변전소의 현황 및 저장 정보를 나타낸다. [표 4]는 기상 관측소의 정보를 나타낸 것으로 부산은 종관 기상 관측 장비 1대와 방재 관측 장비 13대에 대한 정보를 저장하고 있음을 알 수 있다. [표 5]는 기상 관측소 분류에 따른 저장 정보를 나타낸 것으로 종관 기상 관측 장비는 방재 관측 장비와 달리 습도, 기압, 일조량, 이슬점과 같은 기상 상의 정보를 확인할 수 있다. 이에 반해, 방재 관측 장비는 각 소재지의 기온, 강우량, 풍속, 풍향과 정보만 제공한다.

표 4. 부산 내 기상 관측소 기본 정보 및 소재지

구분	코드	이름	소재지	데이터 기간
종관	159	부산청	중구	2005~2014
	160	부산(레)	서구	2005~2014
	910	영도	영도구	2005~2014
방재	921	가덕도	강서구	2005~2014
	923	기장	기장군	2005~2014
	937	해운대	해운대구	2005~2014
	938	부산진	부산진구	2005~2014
	939	금정구	금정구	2005~2014
	940	동래	동래구	2005~2014
	941	북구	북구	2005~2014
	942	대연	남구	2005~2014
	950	사하	사하구	2005~2014
	968	남항	영도구	2005~2014
	969	북항	남구	2005~2014

표 5. 기상 관측소 분류에 따른 저장 정보

구분	기온	강우	풍속	풍향	습도	기압	일조	일사
종관	○	○	○	○	○	○	○	○
방재	○	○	○	○				

[표 6]은 변전소의 전력 데이터의 현황을 나타내며, 전력 데이터는 시간당 전력 수요 정보가 저장되어 있다. 변전소의 위치 정보는 검색할 수 없으므로, 관측소 소재지를 기준으로 분류하였다. 각 변전소의 전력 수요 패턴은 ‘주거, 공단, 특이’세 가지로 분류하였으며, ‘특이’는 다른 패턴과 유사성이 없음을 의미한다.

본 논문에서는 공장지대 및 주거지역에 대한 예측 결과를 비교하기 위해 각 패턴 모형이 나타나는 사하구 지역과 동래 지역의 전력 및 기상 데이터를 활용하여 수요 예측 실험을 진행하였다.

표 6. 부산 내 변전소 현황 및 소재지. [avg(P) : 평균 전력 수요, rank(P) : 평균 전력 수요에 따른 전력 수요 순위]

전력지사	변전소	소재지	구분	avg(P)	rank(P)	
동부산	구서	금정구	주거	65.16	9	
	기장	기장군	특이	43.59	28	
	노포	금정구	주거	12.01	36	
	동래	동래구	주거	64.73	10	
	명례	기장군	특이	4.55	38	
	명장	동래구	주거	83.46	1	
	수영	수영구	주거	72.83	7	
	신양산	기장군	주거	27.94	32	
	용호	남구	주거	50.77	25	
	우동	해운대구	주거	57.47	17	
	정관	기장군	공단	32.16	30	
	좌동	해운대구	주거	56.52	19	
	부산 울산 직할	개금	부산진구	공단	59.28	14
대연		남구	주거	58.86	16	
범일		동구	특이	12.59	35	
범천		부산진구	주거	43.81	27	
부민		서구	주거	53.42	23	
부산		부산진구	주거	54.90	21	
서면		부산진구	특이	73.14	6	
수정		동구	주거	56.12	20	
양정		부산진구	주거	82.68	2	
연산		연제구	주거	27.10	33	
영도		영도구	특이	57.04	18	
북부산		구포	북구	주거	62.09	12
		덕포	사상구	특이	28.31	31
	모라	사상구	공단	73.24	5	
	미남	동래구	주거	48.65	26	
	북부산	사상구	주거	52.38	24	
	엄궁	사상구	공단	63.43	11	
	학장	사상구	공단	59.05	15	
	녹공	강서구	공단	78.79	3	
	녹산	강서구	공단	72.61	8	
	다대	사하구	공단	60.00	13	
서부산	명지	강서구	공단	22.97	34	
	미음	강서구	공단	8.78	37	
	서부산	사상구	특이	53.47	22	
	신평	사하구	공단	78.54	4	
	장림	사하구	공단	33.15	29	

2. 비교 전력 수요 예측 기법

본 논문에서는 기상 정보를 반영한 전력 수요 예측을 시행하며, 이를 검증하기 위해 3가지 비교 실험을 진행하였다. 이는 유사 기온 검색에 따른 전력 패턴 추출 및 신뢰 구간 계산에 따른 전력 예측 기법과 단일 전력

SVR 예측 실험 및 전력/기온 데이터를 사용한 SVR 예측 실험이다.

유사 기온에 따른 전력 패턴 선정 방식의 경우 기온 데이터를 기준으로 일일 최고/최저 기온을 계산하여, 검색 타입 및 허용 범위 구간 값에 따른 유사 기간을 선정하는 것을 말한다. [그림 9]는 검색 타입에 따른 유사 기온 데이터 선별 과정을 나타낸다. 검색 타입이 1이면 최저 기온에 대한 허용 범위 기준을 만족하는 데이터를 학습한다. 검색 타입이 2이면 최고 기온 범위를 만족하는 데이터를, 검색 타입이 0이면 최고/최저 기온 범위를 모두 만족하는 데이터를 선별한다.

검색 결과를 통해 유사 기온이 나타나는 기간 목록이 얻어지면, 해당 일자의 전력 패턴을 학습데이터로 사용한다. 이때, 학습 데이터에 대한 필터링 및 신뢰구간 계산 과정이 수행된다. 인접 시계열 간의 거리가 멀 경우 필터링 되며, [그림 10]과 같이 예외적인 전력 패턴이 제거된다. 유사 전력 패턴을 기준으로 신뢰 구간 계산에 따른 예측 값을 산정하여, 평가 데이터에서 나타나는 전력 예측 값이 신뢰 구간 내에 존재할 경우 정답으로 평가한다. 예측 실험에서 사용되는 신뢰구간은 선정된 표본 데이터에 대한 95% 신뢰구간을 계산하여 구간 안에 있는 경우 예측된 것으로 판단한다.

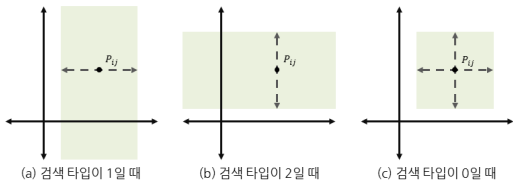


그림 9. 유사 기온 검색 타입에 따른 허용 범위 구간 생성 예시. (a) 검색 타입 1 = 최저 기온, (b) 타입 2 = 최고 기온, (c) 타입 0 = 최고/최저 기온 모두 만족

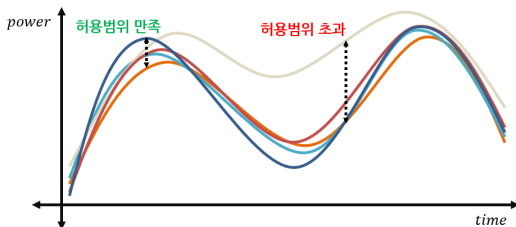


그림 10. 유사 기온 검색 결과로 얻은 전력 데이터에 대한 허용 범위 필터링 과정

유사 기온 데이터에 따른 전력 예측 실험과는 다르게 SVR을 이용할 경우, 기존 연구와 마찬가지로 이전 시간의 전력량을 입력 값으로 한 전력 예측 결과와 k 길이의 이전 전력 수요 패턴을 학습하였을 때의 전력 예측 결과를 비교한다. 이때, 입력하는 k 길이의 전력 수요 패턴에 대해 길이를 조정하면서 예측결과 성능을 평가한다. 기존 연구에서 진행된 회귀 분석 및 기계 학습을 이용한 전력 수요 예측 기법들은 유사 기온에 따른 데이터 필터링을 진행하거나, 학습 데이터가 이전 시간의 전력 수요를 고려하였다. 하지만 본 논문에서는 k 길이에 따른 전력 및 기온 벡터를 학습하여 예측 정확도를 개선하고자 한다.

3. 유사 기온 데이터에 따른 전력 예측 실험

유사 기온 데이터를 이용하여 신뢰구간 계산에 따른 전력 수요 예측 결과는 [표 5]를 통해 확인할 수 있다. [표 7]은 사하 기상 관측 데이터를 기준으로 다대 변전소에 대한 전력 수요 예측을 시행한 결과를 나타낸다. 유사 기온 탐색 타입은 0으로 설정하였으며, 허용 범위 값에 따른 비교 데이터의 증가를 확인할 수 있다. 그러나 정확도의 경우 유사 기온 허용 범위가 6 이상이면 감소하였다. 이를 통해 기온의 단편적인 정보를 패턴 추출에 이용하는 것은 예측 결과에 도움이 되지 않는 것을 확인하였다. 유사 기온 데이터를 추출할 경우 비교 데이터는 충분히 사용할 수 있지만, 기본 예측 정확도가 30~50%로 다른 기법의 실험 결과와 비교하였을 때 활용 가능성이 작으므로 전력 수요 예측에 활용하기 어렵다.

표 7. 다대 지역의 유사 기온 데이터에 대한 전력 예측 결과

유사 기온 범위	평가 일수	평균 비교 데이터	정확도(%)
1	123	2.54	39.66
2	295	5.73	51.22
4	355	16.44	52.25
6	358	28.86	47.27

4. 단일 전력 데이터를 이용한 전력 예측 실험

단일 전력 데이터를 활용하여 전력 예측을 시행한 결과는 [표 8] 및 [표 9]와 같다. SVR을 이용한 전력 예측

의 학습 데이터는 2011년도부터 2013년도까지 시간 단위 전력 수요 데이터를 이용하였다. 평가 데이터는 실제 2014년도의 실제 시간별 전력 수요 데이터를 이용하였다. [표 8]은 신평 지역을 기준으로 이전 전력 데이터를 학습 하였을 때의 전력 수요 예측 결과이며, [표 9]는 동래 지역에 대한 전력 수요 예측 결과를 나타낸 것이다.

k 의 길이가 1인 경우 SVR의 기본 추정 함수와 같이 2차원 좌표계로 치환된 수요 예측 결과를 의미하며, SVM의 학습 결과와 동일한 추정이 가능하다. k 가 1일 때의 수요 예측 결과를 통해 평균 오차율은 낮게 나타나지만, 최대 오차율이 크게 나타나는 것을 확인할 수 있다. k 값 증가하여 전력 벡터에 대한 학습 및 예측 결과를 통해, 전력 수요와 무관하게 오차율이 감소하는 것을 확인할 수 있다. 유사 기온 데이터를 활용하여 신뢰 구간을 계산하는 것과 달리 전력 수요 예측 결과의 평균 오차율이 1~5% 이하로 나타나며, 학습되는 벡터 수가 커질수록 오차율이 감소하는 것을 확인하였다. 또한, 벡터 크기에 따라 오차율 분산과 최대 오차율도 감소하는 것을 확인하였다.

표 8. 신평 지역의 전력 수요 패턴을 학습하였을 때의 SVR 전력 수요 예측 결과 [avg.(Err) : 평균 오차율, var.(Err) : 오차율 분산, max(Err) : 최대 오차율]

k	$pv_i(k)$	$tv_i(k)$	avg.(Err)	var.(Err)	max(Err)
1	○	-	5.0878	28.4757	73.5543
2	○	-	3.7929	16.3735	44.2389
6	○	-	3.2236	11.0759	36.7792
12	○	-	2.8701	8.1112	29.7321
24	○	-	2.5821	5.6191	28.6879

표 9. 동래 지역에 대한 전력 수요 패턴을 학습하였을 때의 SVR 전력 수요 예측 결과

k	$pv_i(k)$	$tv_i(k)$	avg.(Err)	var.(Err)	max(Err)
1	○	-	4.6721	15.9308	68.2423
2	○	-	2.2670	3.6673	47.0927
6	○	-	1.8120	3.0531	37.0804
12	○	-	1.5349	2.0775	28.4417
24	○	-	1.2943	1.5730	26.3768

[표 8]과 [표 9]의 경우, 서로 다른 전력 수요 패턴을 가지는 신평(공장 지역 수요 패턴)과 동래(주거 지역 수요 패턴)에 대한 비교 실험 결과 SVR을 적용하였을

때 오차율 범위가 0.4~1.3% 정도로 차이가 크지 않으나, 동래 지역 예측 결과의 분산 및 최대 오차율이 더 낮은 것을 확인하였다. 이는 한 주의 수요 패턴 모형이 일정하여 주말에 대한 예측 오차가 작게 나타났기 때문이다.

5. 기온 및 전력 패턴을 이용한 전력 예측 실험

전력 및 기온 패턴을 이용하여 전력 예측을 시행한 결과는 [표 10], [표 11]과 같다. 기온 데이터를 학습에 이용할 경우, 일일 최고/최저 기온 정보만 활용하는 것이 아닌 k 길이만큼의 이전 기온 패턴 정보를 함께 사용한다. [표 10]은 신평 지역에 대해 기온 및 전력 데이터를 학습하여 전력 수요 예측을 시행한 결과이다. 이때, 단일 전력데이터를 활용하는 것과 오차율 차이는 0.1~0.3 사이로 근소한 차이를 보인다. 하지만 최대 오차율이 16~35%로 낮게 나타나며, 단일 전력 수요 패턴을 이용하는 것보다 평균 19%의 최대 오차율 감소를 확인할 수 있다.

[표 11]은 동래 지역에 대한 전력 및 온도 데이터를 이용한 전력 수요 예측 결과이다. 신평 지역과 마찬가지로 단일 전력 데이터를 학습하는 것과 평균 오차율 차이는 근소하게 나타나지만, 최대 오차율의 감소가 평균 16%가 되는 것을 확인하였다.

표 10. 신평 지역에 대한 전력 및 기온 패턴을 학습 데이터로 주었을 때의 SVR 전력 수요 예측 결과

k	$pv_i(k)$	$tv_i(k)$	avg.(Err)	var.(Err)	max(Err)
1	○	○	5.1068	28.4344	35.2501
2	○	○	3.9890	16.1360	23.7692
6	○	○	3.3764	10.7564	21.6678
12	○	○	2.9614	7.5473	20.6893
24	○	○	2.5127	4.6958	16.3512

표 11. 동래 지역에 대한 전력 및 기온 패턴을 학습 데이터로 주었을 때의 SVR 전력 수요 예측 결과

k	$pv_i(k)$	$tv_i(k)$	avg.(Err)	var.(Err)	max(Err)
1	○	○	4.6289	15.5845	31.3908
2	○	○	2.1307	3.2626	26.3461
6	○	○	1.7821	2.3911	25.3235
12	○	○	1.6767	1.7324	21.9117
24	○	○	1.6929	1.5136	18.0308

전력 및 기온 패턴 데이터를 활용하여 전력 수요 예측 실험을 진행하였을 때, 단일 전력 데이터를 활용하는 것의 평균 오차율 차이는 거의 없는 것을 알 수 있다. 하지만 단일 전력을 이용하여 수요 예측을 시행한 것보다 최대 오차율의 차이 평균 15% 이상 개선되는 것을 확인 가능하다. 이를 통해, SVR을 활용하는 데 있어서 전력과 기상 패턴 정보를 함께 사용하는 것이 예측 오차율을 개선하는 데 도움이 되는 것을 알 수 있다. 또한, 주어지는 패턴 길이 k 가 짧을 경우에도 최대 오차율이 낮은 것을 통해 학습 데이터가 충분하지 않을 때도 예측 정확성이 높게 나타난다. 하지만, 단일 전력 데이터를 활용하는 것보다 학습에 필요한 데이터가 방대하여 수행시간이 길어지는 문제가 발생하며, 오차율 개선이 이루어지지 않는 경우도 존재한다.

V. 결론 및 향후 연구과제

본 논문은 전력 수요 예측에 대해 국내외 실증 사례 및 현재 사용되고 있는 전력 수요 예측 기법을 분석하였다. 또한, 회귀 함수가 적용된 기계학습 기법인 SVR을 이용하여 전력 및 기상 데이터를 동시에 고려하였을 때의 전력 수요 예측 실험을 진행하였다. 이를 검증하기 위해 단일 전력 및 유사 기온에 대한 전력 수요 예측 비교 실험을 진행하여, 기상 정보를 활용하였을 때의 예측 정확도를 비교하여, 전력 데이터만을 사용한 수요 예측 결과와의 최대 오차율 차이가 15% 이상 개선되는 것을 알 수 있다.

실제 변전소 데이터를 활용하는 과정에 있어서, 데이터 한계로 인해 전력 데이터 이외에 기온 데이터를 이용한 실험을 진행하였다. 이는 강수량과 천재지변과 같은 특수 상황을 반영하기에는 데이터양이 풍부하지 않고 2011년에서 2014년까지의 기후 조건 가운데 기온 데이터만 활용 가능하였다. 또한, 지역적 특성 및 부산 지역에 대한 실험을 진행하지 않았기 때문에 오차율에 대한 신뢰도를 만족하지 못한다. 이를 개선하기 위해 다음과 같은 추가 구현 및 실험을 진행하고자 한다.

전력 시계열 데이터에 대해 SVR이 아닌 비선형 회귀 분석 모델 및 전력 시계열 패턴에 대한 분류를 적용하

여 추가적인 실험을 시행한다. 또한, 전력 및 기상 데이터를 추가 확보하여, 특수 상황 및 강수량 정보와 같이 본 논문에서 활용하지 못한 데이터를 추가로 적용하여 다중 특성에 대한 전력 수요 예측 기법을 개선한다. 이외에 PJM과 같이 여러 기법을 순차적으로 적용하거나 피드백 과정을 거쳐 오차율을 감소시키는 방안 대해서도 고려하고자 한다. 이를 위해 기존 전력 수요 예측 연구에서 나타나는 회귀 분석 모델[13][14]과 다중 시계열에 대한 엔트로피 연구를 응용하고자 한다[15].

참고 문헌

- [1] M. Lee and S. Kim, "Analysis of smart grid technologies and domestic and foreign policy trends," Journal of Korean Institute of Information Technology, Vol.11, No.8, pp.181-187, 2013.
- [2] 고종민, 송재주, 김영일, 양일권, "전력소비자의 단기수요예측을 위한 전력소비패턴과 환경요인과의 관계 분석," 전기학회논문지, Vol.59, No.11, pp.1956-1963, 2010.
- [3] K. S. G. Association, "Smart grid ami(advanced metering infrastructure)," 2012.
- [4] D. Basak, S. Pal, and D. C. Patranabis, "Support vector regression," Neural Information Processing Letters and Reviews, Vol.11, No.10, pp.203-224, 2007.
- [5] Madasu Hanmandlu and Bhavesh Kumar Chauhan, "Load Forecasting Using Hybrid Models," IEEE Transactions on Power Systems, Vol.26, No.1, 2010.
- [6] Ying Chen, Peter B. Luh, and Stephen J. Rourke, "Short-Term Load Forecasting: Similar Day-Based Wavelet Neural Networks," IEEE Transactions on Power Systems, Vol.25, No.2, 2010.
- [7] James W. Taylor, "Short-Term Load Forecasting With Exponentially Weighted Methods," IEEE Transactions on Power Systems, Vol.27, No.1, 2012.
- [8] Capacity Adequacy Planning Department, "PJM Load/Energy Forecasting Model," PJM, 2007.

[9] "PJM Manual 11: Energy & Ancillary Services Market Operations," PJM, 2012.

[10] Dennis Gaushell, "Overview of Electric Load Forecasting at CAISO," CAISO, 2007.

[11] 한국전력거래소, "전력과 기상 특성 분석을 통한 차기 전력수요예측 시스템 개발전략 수립에 관한 연구," 2014(3).

[12] 한국전력거래소, "실시간 수요예측 기법 개발 및 온라인 수요예측 방안 연구," 2015(2).

[13] 남봉우, 송경빈, 김규호, 차준민, "다중회귀분석법을 이용한 지역전력수요예측 알고리즘," 조명·전기설비학회논문지, Vol.22, No.2, pp.63-70, 2008.

[14] 송경빈, "시간대별 기온을 이용한 전력수요예측 알고리즘 개발," 전기학회논문지, Vol.63, No.4, pp.451-454, 2014.

[15] M. Božić, M. Stojanović, Z. Stajić, and N. Floranović, "Mutual information-based inputs selection for electric load time series forecasting," Entropy, Vol.15, No.3, pp.926-942, 2013.

저자 소개

탁 해 성(Haesung Tak)

정회원



- 2013년 : 부산대학교 정보컴퓨터공학부(공학사)
- 2015년 : 부산대학교 전기전자컴퓨터공학과(공학석사)
- 2015년 9월~현재 : 부산대학교 전기전자컴퓨터공학과 박사과정

<관심분야> : 데이터마이닝, 기계학습

김 태 용(Taeyong Kim)

준회원



- 2015년 : 부산외국어대학교 임베디드IT과(공학사)
- 2015년 3월~현재 : 부산대학교 전기전자컴퓨터공학과 석사과정 관심분야

<관심분야> : 데이터마이닝, 기계학습

조 환 규(Hwan-Gue Cho)

정회원



- 1984년 : 서울대학교 계산통계학과(이학사)
- 1986년 : KAIST 대학원 전산학과(공학석사)
- 1990년 : KAIST 대학원 전산학과(공학박사)

- 1990년 3월~현재 : 부산대학교 컴퓨터공학과 교수

<관심분야> : 계산이론, 생물정보학

김 희 제(Heeje Kim)

정회원



- 1980년 : 부산대학교 전기공학과(공학사)
- 1982년 : 부산대학교 전기공학과(공학석사)
- 1990년 : 규슈대학교 대학원 에너지공학과(공학박사)

- 1995년 11월~현재 : 부산대학교 전기공학과 교수

<관심분야> : 레이저, 플라즈마, 태양전지 공학